IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 3, November 2024

Natural Language Processing for Electronic Health Record Optimization in Android Applications

Gift Mboneche[1] and Ulemu Mponela[2]
Student, DMI St John The Baptist University, Lilongwe, Malawi[1]
Lecture II, Department of Computer Science, DMI St John The Baptist University, Lilongwe, Malawi[2]
mbonechegift@gmail.com and Ulemu44@gmail.com

Abstract: *This paper aims to demonstrate the development of an Android application designed to improve healthcare data management using Natural Language Processing (NLP) technologies. It develops an Android application that enables healthcare professionals to access patient records anytime, enhancing decision-making. The system uses Optical Character Recognition (OCR) to extract patient data from documents, simplifying data entry for diagnoses and treatments. It also incorporates Quick Response (QR) code generation for accurate data retrieval. By leveraging NLP, the mobile application will enhance efficiency, accessibility, and accuracy in healthcare, particularly benefiting developing countries where digital health records are uncommon..*

Keywords: Natural Language processing, Optical Character Recognition, Quick Response, Electronic Health Records

I. INTRODUCTION

Natural Language Processing (NLP) is gaining prominence in healthcare due to its potential to enhance Electronic Health Records (EHR) systems by addressing inefficiencies in data management. With the vast amount of unstructured healthcare data, such as clinical notes and medical reports, traditional data entry and retrieval methods are becoming increasingly burdensome. NLP techniques, including Optical Character Recognition (OCR), allow for the automatic extraction of key patient information from printed and digital documents, improving the accuracy and efficiency of healthcare data management. These technologies enhance clinical decision-making and patient care enabling mobile access to patient records and incorporating QR code generation for streamlined data retrieval. This project aims to utilize NLP in developing an Android application that integrates such features, promoting the adoption of EHRs, especially in developing countries where such technologies are underutilized. Despite the challenges related to data privacy and interoperability, the advancement of NLP holds significant promise for transforming healthcare delivery.

II. OBJECTIVES

The objectives of the project "Natural Language Processing for Electronic Health Record (EHR) Optimization in Android Applications" are focused on addressing the healthcare challenges faced in developing countries like Malawi, where physical health passports are still widely used. This paper first aims to demonstrate the development of an efficient, user-friendly mobile application imitating an EHR system by leveraging Natural Language Processing (NLP) techniques to digitize patient data and optimize its management. By incorporating the extraction and analysis of clinical information from unstructured text, such as handwritten medical records, the project seeks to streamline consultation and improve healthcare professionals' access to patient information.

Another key objective is to enhance data security and privacy measures, ensuring that sensitive patient information is protected from unauthorized access. The project also aims to promote the use of current and emerging technologies in handling patients' data by integrating data storage and retrieval enabling seamless communication between healthcare personnel and the community of clients particularly in developing countries. Additionally, it seeks to empower healthcare professionals with real-time decision support tools, allowing for better-informed clinical decisions based on

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 3, November 2024

comprehensive patient data. Ultimately, these objectives aim to overcome the limitations of manual data entry, fragmented healthcare systems, and restricted access to patient information, improving the efficiency, quality, and accessibility of healthcare services in developing countries.

III. EXISTING SYSTEMS VS PROPOSED SYSTEM

Several existing systems have emerged inrecent years that aim to improve healthcare data management.Many havenot provided comprehensive, mobile-accessible solutions tailored to low-resource settings. Epic is a leading Electronic Health Record (EHR) system that offers robust health services for patient data management and clinical documentation. Epic has introduced mobile capabilities through its MyChart app, allowing patients to access their medical records. Still, its primary focus remains on large healthcare organizations in developed countries, making it costly and complex to use in settings like Malawi where infrastructure is limited. Cerner continues to be another significant player in the EHR space, offering advanced features like clinical decision support and patient data analytics. In 2020, Cerner introduced some Natural Language Processing (NLP) capabilities to improve data extraction from clinical notes, enhancing the accuracy of EHR systems. However, while Cerner has made strides in integrating mobile features, its reliance on cloud infrastructure and advanced technologies limits its applicability in environments with inconsistent connectivity.

The Open-source platform has been a game-changer in resource-constrained settings. Updated regularly, it allows for basic patient record management and is customizable based on local needs. However, its functionality is still relatively basic, and it lacks advanced features like Optical Character Recognition (OCR) for extracting patient data from printed documents. While it has mobile extensions, these are often limited to data collection and do not fully support the needs of healthcare professionals seeking real-time, comprehensive patient information.

In addition, OpenMRS is a powerful tool in low-resource settings, mainly focused on public health data collection and reporting. Recent updates have enhanced its mobile functionalities, allowing offline data collection and reporting. However, its core design is not suited for individual patient care or detailed medical consultations, which are crucial for healthcare delivery in clinical settings.

While these systems have madenotable contributions to healthcare data management, none are specifically designed to provide the mobile, real-time access to comprehensive patient information needed to streamline medical consultations. Moreover, they often lack integration of NLP and OCR technologies to digitize paper-based records, which are still prevalent in many developing regions. The proposed mobile application addresses these gaps by enabling healthcare professionals to access patient information anytime, anywhere, with automated data extraction from digital and printed records. This system will offer a cost-effective and scalable solution that supports real-time decision-making in low-resource settings like Malawi, where current systems have limited reach.

IV. SYSTEM DESIGN AND IMPLEMENTATION

The design and implementation of the mobile application for optimizing Electronic Health Records (EHR) is centered around enhancing healthcare delivery through real-time access to comprehensive patient data. The system is developed using Android Studio, the official Integrated Development Environment (IDE) for Android development. This platform provides robust tools for creating mobile applications and supports integratingtechnologies such as Optical Character Recognition (OCR) and Natural Language Processing (NLP). Below is an in-depth look into the system design process, the hardware and software requirements are the building blocks that facilitate the development process of the mobile application.

4.1 Design Process

The design process followed a modular approach to ensure the system could integrate different features, including patient data management, real-time access to health records, and OCR for digitizing printed documents. The primary objective was to ensure that healthcare professionals could quickly and securely retrieve patient information, even in environments with limited resources, and to have access to critical health information for patients that might otherwise be inaccessible due to the loss of physical health passports.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 3, November 2024

4.1 Hardware and Software Requirements

Hardware Requirements:

- *Android Device:* A device running Android 6.0 (Marshmallow) or higher is required for the application to function optimally.
- *Minimum RAM*: 2 GB
- *Processor:* ARM-based processor
- *Storage:* At least 100 MB of free storage for installation and record storage.
- *Camera:* A camera for scanning QR codes and capturing images for OCR processing.

Software Requirements

- *Android Studio:* Version 4.1 or later for development, running on Windows, macOS, or Linux.
- *Firebase SDK:* For authentication, database management, and cloud services.
- *Google MLkit text-recognition and MLkit Barcode Generation APIs*: Used for integrating Optical Character Recognition (OCR) functionality and QR Code generation.
- *Bumptech Glide API:* Used for rendering images in the application.
- *Gradle Build System:* To manage project dependencies and build processes.
- *Java:* Primary programming language used for the development of the application.
- *Minimum Android OS Version:* Android 6.0 (Marshmallow) or above.

The system's key modules include:

- *Authentication and User Management:* Ensuring secure access to the system for healthcare providers using Firebase Authentication.
- *Patient Data Handling*: Digitizing and storing patient records using Firebase Realtime Database, allowing easy access and updating of information.
- *Optical Processing:* Leveraging Google's OCR technology to extract patient details fromprinted or un-editable (pdf) records and NLP language models to organize and easily classify the language being extracted.
- *QR Code Generation:* Enabling secure and accurate retrieval of patient data during consultations or medical procedures.
- *User Interface:*The user interface (UI) was designed to be intuitive, offering easy navigation and quick access to patient records, diagnoses, and treatment plans.
- *Account password reset module:* This module facilitates the reset of passwords and account restoration for authentication purposes as well as security to user accounts, an email is verified first against the user who requests a password reset. Once the email is verified, a verification code is sent to the user for further verification, upon success, a reset link is sent to the email, finally resetting the account's password, enhancing security and privacy.

V. METHODOLOGY

The system development process emphasizes iterative development and feedback-driven refinement, allowing continuous improvement and adaptation based on user feedback and evolving requirements. This agile approach enables rapid prototyping, testing, and iteration, ensuring that the mobile app meets the expectations of healthcare professionals and end-users in consultation scenarios. We adopt an agile approach to development, prioritizing flexibility, and iterative progress. Development occurs in sprints, with frequent releases and feedback loops ensuring alignment with user needs and project goals. This methodology enables rapid adaptation to changing requirements throughout the development process.

5.1 Input Design Screenshots

These Screenshots depict the applications' input channels including user authentication such as a Login page, Account creation page, Personal details input areas and the creation of Quick Response (QR) codes for the user.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 3, November 2024

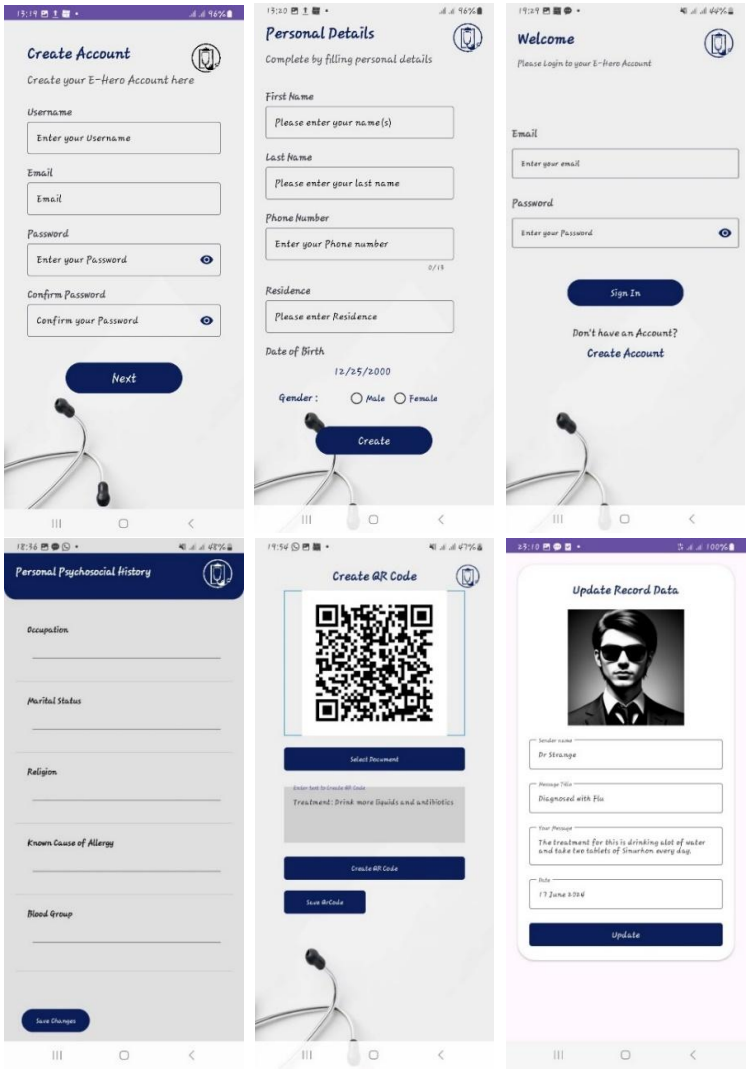

Fig 1. Application Input Screenshots

5.2 Output design Screenshots

These images depict scenarios where the user views records, creates Quick Response (QR) codes, and extracts texts from digital and printed documents using the mobile camera.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 3, November 2024

Fig 2. Application Output Screenshots

5.3 System Architecture

The system follows a client-server architecture, where the Android application serves as the client, and Firebase handles backend processes such as authentication, real-time database operations, and file storage. Data flows from the client (healthcare professional) to the server (Firebase), where patient records are retrieved, updated, and stored.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 3, November 2024

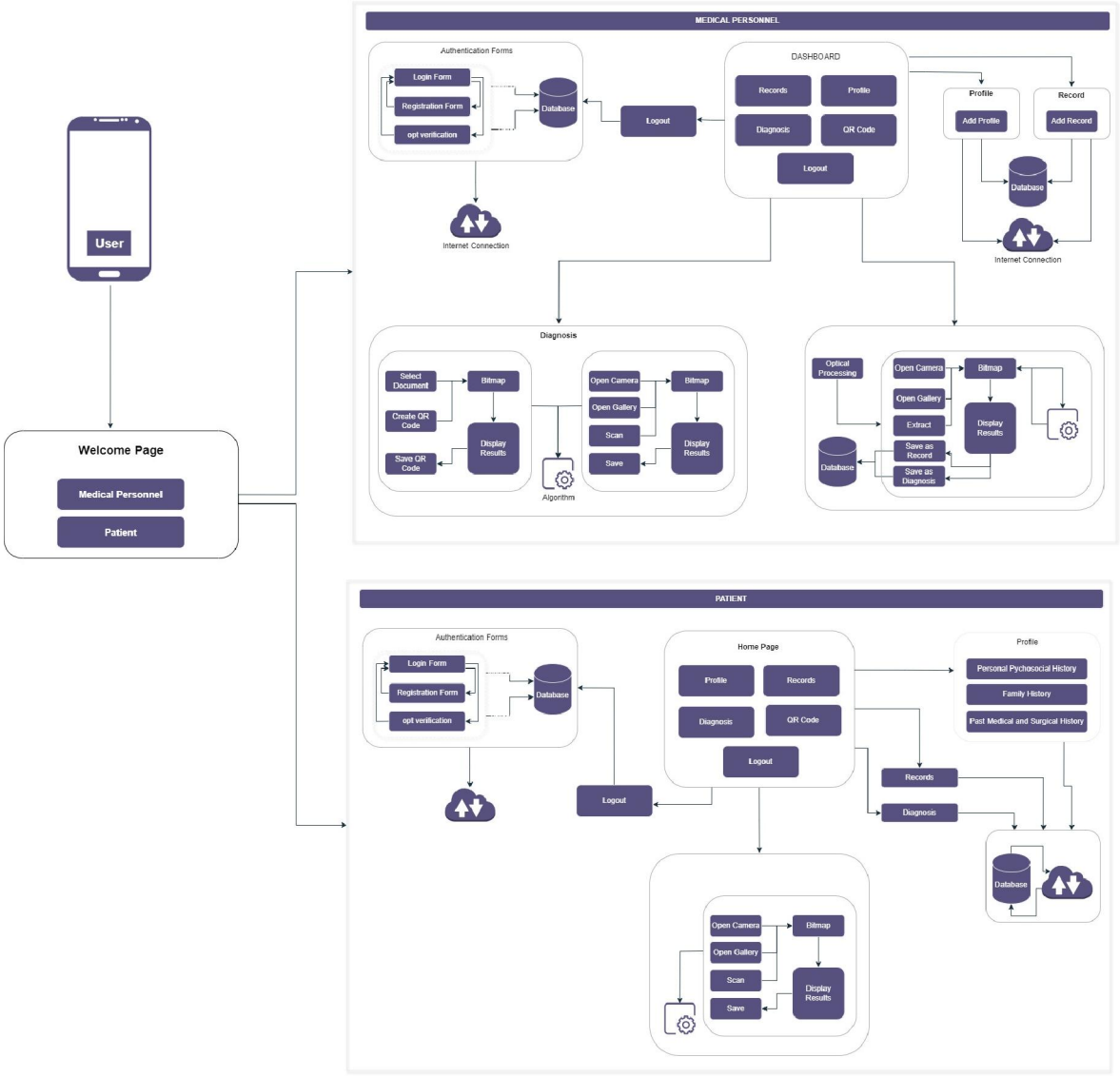

Figure 3. System Architecture

5.4 Data Flow Diagram

This data flow diagram depicts how data will be handled throughout the mobile Application and how several system features will interact with that data from its inception to its output. This will give an overview of how data will be used and manipulated throughout the system. The user node depicted in the diagram, be it a patient or doctor upon user-type selection, enters credentials for login validation, or registers an account which then the data is stored in the database, then the users enter the dashboard where they can access other modules of the application, the medical user in this case can create records for patient, edit, store and send to patients, in creating the record, the medical user can generate QR Code from text, as well as extract text from printed and un-editable digital documents through the Optical Processing module.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 3, November 2024

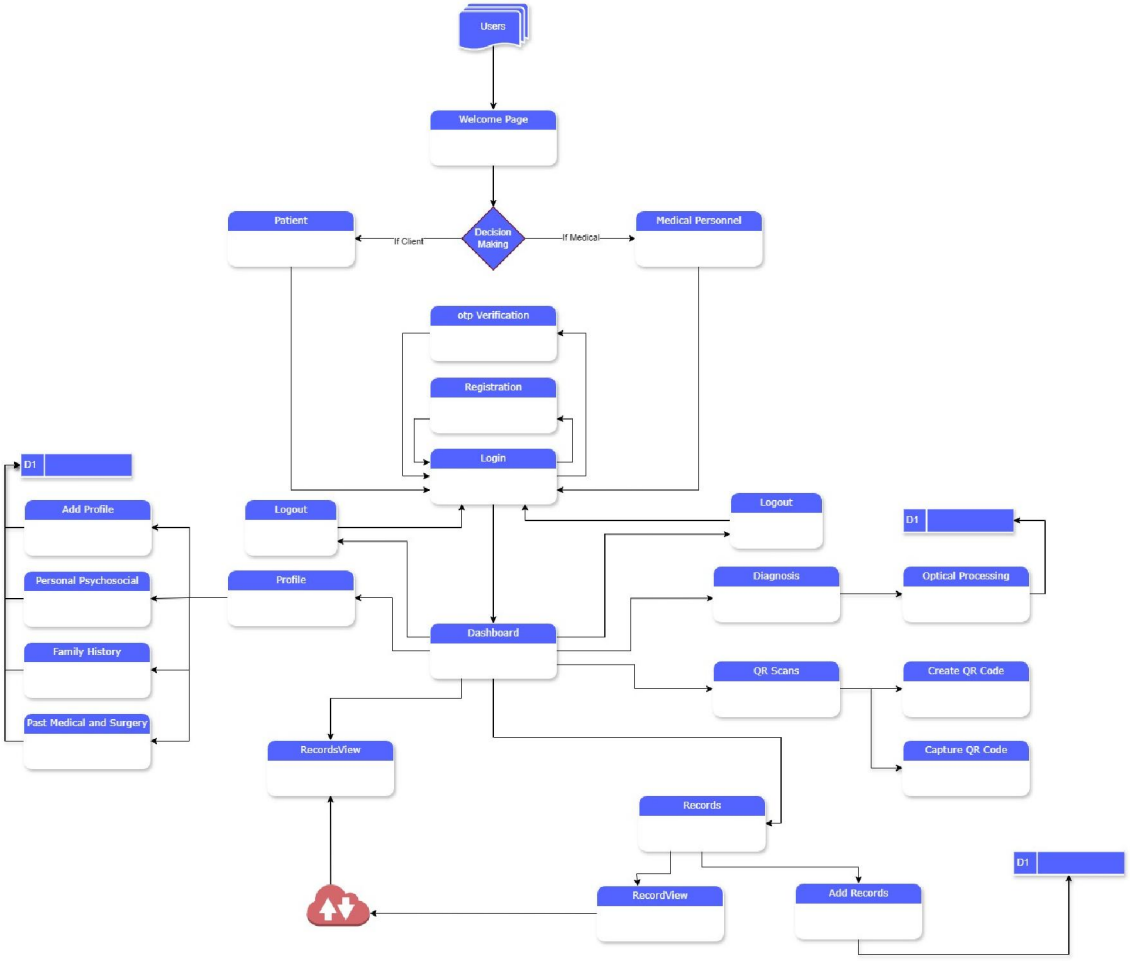

Figure 4 Data Flow Diagram

5.5 Class Diagram

The input diagram illustrates the classes used in the system and their respective attributes and methods that will cater to the application operations, the entry point to the application is the *"MainActivity.class"* which authenticates users to the main dashboard if already signed in to the application. If not authenticated yet, this activity initiates the *"userType.class"* activity that provides an option to the user to select a user type so as to proceed in the application, be it Patient or Medical Personnel. Once the selection is done, the application proceeds to the Login page and Register options in the *"LoginActivity.class"*. The Login Activity leads to the "Home" class after proper and successful login and registration. The "Home" class is the main entry to the core application functionalities providing the user with all use cases found in the application.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 3, November 2024

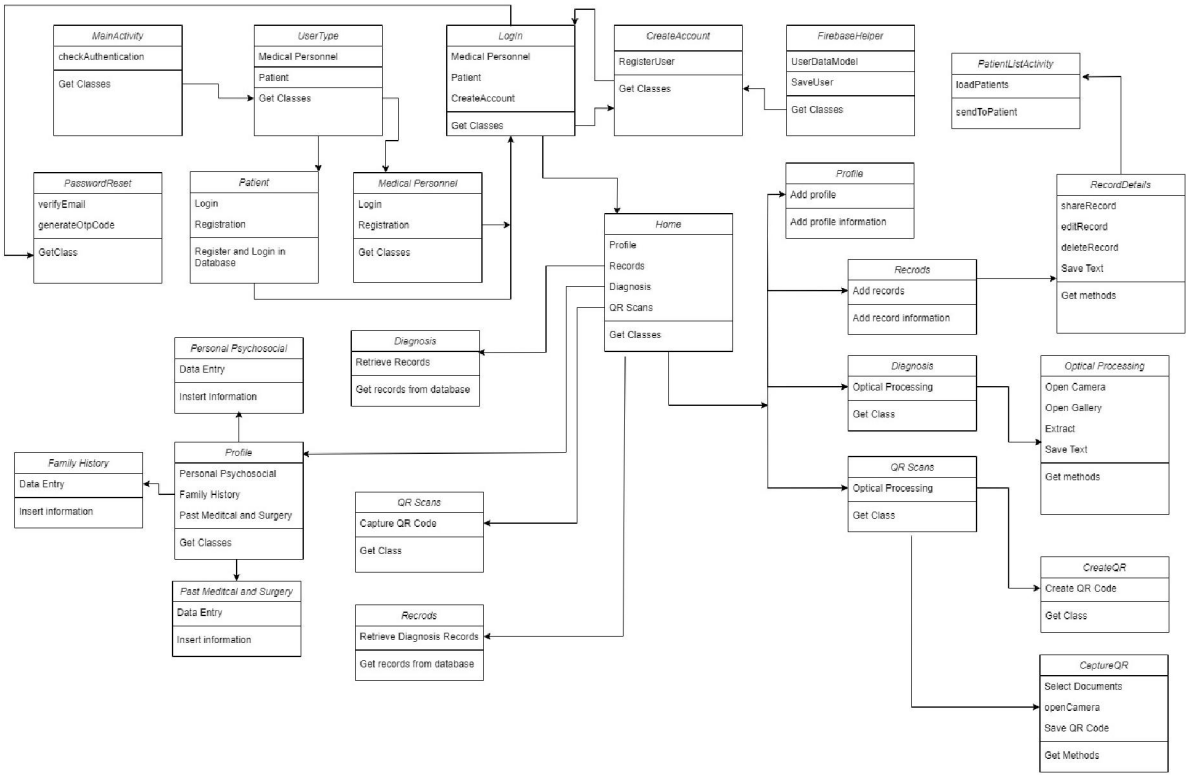

Figure 5. Class Diagram

VI. RESULTS

The various tests the application underwent reported to be an efficient tool as it proved to be helpful in extracting text for data entry purposes and other data manipulation scenarios made possible by the Optical Character Recognition (OCR) feature, the application also proved to be timely in data retrieval when sent to from a medical user to a client/patient due to its integration of Firebase real-time database which allows real-time capabilities. The application's ability to maintain data showcased its importance in consultation scenarios as it proved to have undeniable potential to provide comprehensive information for medical users to efficiently aid them in improving the delivery of health services to patients during consultation processes.

VII. DISCUSSION

This application showcases a high potential to help individuals keep track of their personal information as well as provide comprehensive information to medical users, especially in consultation processes, this can be an essential tool that can provide additional value to the already existing systems in the health sector as its role does not replace the existing systems but rather provide additional efficiency in consultation scenarios which most times are not efficient especially in conditions where the client/patient does not provide a comprehensive medical background or history which can negatively affect the way health services are delivered in that particular moment. Though this is the case, additional research should be considered as consultation scenarios are different in various countries as well as in various health institutions, hence the implementation methods may vary due to these differences.

VIII. CONCLUSION

The healthcare sector in developing nations can greatly benefit from technologies like Natural Language Processing (NLP), QR codes, and the digitization of physical documents. NLP helps automate the understanding and processing of medical information, making it easier for healthcare providers to access and use critical patient data. QR codes ensure

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 3, November 2024

the secure and quick retrieval of medical information, reducing the risk of errors and improving patient safety. Digitizing paper records makes data management more efficient, enhances confidentiality, and allows faster access to patient information. These technologies can transform healthcare systems by improving the accuracy, accessibility, and security of medical records, leading to better healthcare outcomes.

Furthermore, this mobile application showcases the possibility as well as the potential for developing countries having less access to advanced technological infrastructure to equally excel in healthcare service delivery while utilizing current technology available to them which most times are overlooked and underutilized, which if taken advantage of can efficiently optimize healthcare operations, improving how consultations are approached with adequate information, ultimately providing medical professionals with better decision support tools and improved work lifestyle.

IX. ACKNOWLEDGMENT

First and foremost, I am grateful to the almighty God for the strength, health, and above all life. I am thankful for the support from family for their unwavering support. I would like to thank the DMI – St. John the Baptist University Malawi, for providing me the opportunity to do the journal work as part of our curriculum. I also sincerely thank Mr. Ulemu Mponela for his support in making this paper possible as well guiding throughout my curriculum project as well as Mr. Mtende Mkandawire , Head of Computer Science and Information Technology for their kind help during my project work by providing an opportunity to enhance my career.

REFERENCES

[1]. Murphy, K. (2020). Epic Systems' Mobile Health Strategy: EHR on the Go. *Healthcare IT News*. Retrieved from Healthcare IT News

[2]. Smith, T. (2021). Cerner's EHR Solutions and the Evolution of NLP in Healthcare. *Journal of Healthcare Informatics*,7(2), 56-67.

[3]. Johnson, A. (2020). Enhancing Clinical Documentation with NLP: A Look at Cerner's Recent Advances. *Health Tech*, 14(3), 12-19.

[4]. OpenMRS Community. (2022). OpenMRS: Adapting to the Needs of Low-Resource Healthcare. Retrieved from OpenMRS

[5]. Duffy, M. (2021). Open-Source EHR Solutions in Global Health: The Case of OpenMRS. *Global Health Informatics Journal*, 9(1), 28-34.

[6]. DHIS2 Team. (2023). DHIS2 Mobile: Advancing Public Health Data Management. Retrieved fromDHIS2

[7]. Zhang J, Lin H, Yang Z, Lin X, Li X, A Review of Natural Language Processing Approaches to Extract Drug-Related Adverse Effects from Electronic Medical Records, (2019).

[8]. Wang, Y., & Jiang M, Building large-scale, diverse, and balanced clinical corpora from electronic health records: the Shanghai i2b2 corpus, (2020).

[9]. Elbattah M, Fernandez-Luque L, Rasmussen K, Bjørn P, Fadnes LT, A Systematic Review of Approaches for Natural Language Processing of Clinical Narratives, (2021).

[10]. Roberts K, Shooshan SE, Rodriguez L, Abner E, Knoth A, Jackson SA, Natural Language Processing for Electronic Health Record-Based Research: A Systematic Review, (2020).